

## ■参加者紹介

電機メーカーに勤務するテストエンジニア(入社1年目)です。製品に対して手動で行われているテストを、どのようにして自動化するかを検討し、実際に自動化環境を構築する業務を主に担当しております。

## ■はじめに

以降では、聴講した発表の中で特に興味深かったものをピックアップし、各発表の概要と、聴講した感想を述べます。

## ■2017年3月14日(1日目)■

□講演情報      Tue March 14 (Day-1) 9:30 AM

### KEYNOTES

タイトル The State of Continuous Integration Testing at Google.

講演者 John Micco, Google, USA

資料 <http://aster.or.jp/conference/icst2017/program/jmicco-keynote.pdf>

#### [発表の概要]

Google のシニアマネージャーである John Micco 氏による、Google におけるテスト活動についての基調講演が行われました。Google では現在、420 万件ものテストが自動で継続的に実行されており、その膨大な数のテスト実行を上手く管理するためのシステムが運用されているそうです。過去 2 年間のテスト結果はデータベースに蓄積されており、講演では、1 ヶ月間のデータを分析して得られた知見を共有して下さいました。また、Flaky Test と呼ばれる種類のテストについて説明が行われました。Flaky Test とは、結果が一定ではない(Pass になったり Fail になったりする)テストを指し、マルチスレッドによる並列実行や実行環境の違い等の要因によって引き起こされます。原因の究明や修正作業によって開発者の生産性が低下してしまうため、Flaky Test は Google の方々にとって悩みの種となっているそうです。

#### [感想]

Google の方々も、Flaky Test のような再現性の低い問題への対処に苦心しているのだな、と親近感を覚えました(規模感はまるで違いますが、私自身も再現性の低いバグに苦しんだ覚えが多数あります)。Flaky Test を不可避の問題だと想定し、発生した際の対処を工夫する、という姿勢を取ることが大事だと学びました。また、テスト結果が蓄積されたデータベースは様々な研究の種になりそうだ、と期待が膨らみました。今回の基調講演のような、情報共有の場をたくさん設けて頂けると良いなと思います。

---

---

□講演情報      Tue March 14 (Day-1) 11:00AM

## I01: Industry #1

1. タイトル Towards A Testbed for Automotive Cybersecurity

講演者    Session chair: Murat Ozturk

Daniel S. Fowler, Madeline Cheah, Siraj Ahmed Shaikh and Jeremy Bryans

[発表の概要]

自動車に搭載されているコンピュータのセキュリティテストに関する発表でした。自動車に搭載されているコンピュータがハッキングされた場合、ハンドルやブレーキ、クラクションやワイパーに至るまで、ほとんどの装置が遠隔で操作されてしまう恐れがあります。ハッキングは乗車している人の命に関わるため、セキュリティテストが重要視されています。実際に自動車を用いてテストを実施するとコストを要してしまうため、シミュレーション(テストベッド)が積極的に活用されているそうです。

2. タイトル How to Test in Sixteen Languages? Automation Support for Localization Testing

[発表の概要]

製品のローカライズ(特定の国・地域に製品を展開するための調整)が正確に行われているかを確認する、ローカライズテストを自動化する試みについての発表でした。ローカライズテストにおいては、アプリケーションの GUI 上で表示される文言や、その配置に誤りがないかを確認する作業に、時間や人手を要するという問題点が存在します。その問題を解決するため、オリジナル言語(英語)の GUI をリファレンスとして保持し、ローカライズされた GUI をリファレンスと比較することで、文言やその配置の誤りを自動で検出します。

[セッションを通しての感想]

私自身がテストに関する業務に携わっていることもあり、産業界で実際に行われているテストやその改善活動のお話はとても興味深く、刺激になりました。特にローカライズテストに関してはこういったテスト活動が行われていることを知ることができ、勉強になりました。

■2017年3月15日(2日目)■

□講演情報      Tue March 15 (Day-2) 11:00AM

## R04: Regression Testing

タイトル    Perphocy: Performance Regression Test Selection Made Simple but Effective

講演者    Session chair: Shin Yoo

Augusto Born de Oliveira, Sebastian Fischmeister, Amer Diwan,

Matthias Hauswirth and Peter F. Sweeney.

[発表の概要]

パフォーマンスの回帰テストを効率よく選別する方法についての発表でした。回帰テストはデグレードの検出を目的とした重要なテストですが、テスト対象に変更が加えられる度に全てのテストを実施するとコストが大きくなってしまいます。加えた変更がパフォーマンスに影響を及ぼす場合のみテストすることが理想です。この発表では、加えた変更内容(静的なデータ)と過去のベンチマーク実行時のデータ(動的なデータ)を分析することで、変更がパフォーマンスに影響を及ぼすかどうかを予測する手法が示されました。

---

---

[感想]

回帰テストはテスト自動化の主な対象となるため、本発表の内容は大変参考になりました。過去に携わった業務では、自動化対象とした回帰テストの規模はそこまで大きくなかったため、テスト対象に変更が加わるたびに自動テストを実施していました。しかし、テスト結果の分析は人手で行う必要があり、実施する度に分析のコストが嵩みます。テスト実施機会の選別を行うことで、分析のコストを削減することができるため、本発表のアプローチを参考に、テスト実施機会の選別が可能かどうかを検討する意識を持ちたいと思います。

□講演情報      Tue March 15 (Day-2) 14:00PM

## **R05: Web and Mobile Applications**

タイトル ATOM: Automatic Maintenance of GUI Test Scripts for Evolving Mobile Applications

講演者    Session chair: Nan Li

Xiao Li, Nana Chang, Yan Wang, Haohua Huang, Yu Pei, Linzhang Wang and Xuandong Li.

[発表の概要]

GUI テストスクリプトのメンテナンスを自動化する手法についての発表でした。アプリケーションの GUI に変更が加わると、過去に作成された GUI テスト用スクリプトにも変更を加える必要があります。モバイルアプリケーションにおいては GUI の変更が頻繁に行われるため、既存スクリプトのメンテナンスに要するコストが問題視されています。この発表では、アプリケーションのモデルを入力として与えることで、GUI に加わった変更を反映して既存スクリプトをメンテナンスする手法について示されました。現状、モデルの作成を手動で行う必要があるようですが、モデル作成に要するコストを削減する、ひいてはモデル作成自体を自動化する手法についても研究が進められているとのことでした。

[感想]

本発表で扱われている課題は私も以前から認識しており、その課題の解決を目的とする研究分野には以前から期待を寄せていましたが、本発表を聴講したことでその期待が更に大きくなりました。アプリケーションの GUI が複雑になるに従って、モデル作成に要するコストも増大するため、モデル作成を補助する手法の発展を望みます。

---

---

■2017年3月16日(3日目)■

□講演情報 Tue March 16 (Day-3) 11:00AM

## R07: Empirics on Testing

タイトル Broadcast vs. Unicast Review Technology: Does it Matter?

講演者 Session chair: Mike Papadakis

Armstrong Tita Foundjem, Foutse Khomh and Bram Adams.

### [発表の概要]

コードレビューの形式とレビューの質との関係性についての発表でした。Broadcastとは、あるコミュニティに所属する不特定多数の人によってレビューを受ける形式を指します。また、Unicastとは、JIRA等のバグ管理システム上で特定の人を指定してレビューを受ける形式を指します。レビュー形式をBroadcastからUnicastに移行したプロジェクトを対象に、レビューの質を定量的に評価する尺度を定め調査したところ、2つのレビュー形式の質に有意な差があることが判明したそうです。具体的には、UnicastはBroadcastと比較してレビューに2倍の時間を要しますが、バグを見逃してしまう割合は少なくなることがわかりました。

### [感想]

レビューの質を定量的に評価する研究分野の存在を知らなかったため、本発表は大きな学びとなりました。また、調査結果は直感に合っており、納得できるものでした。

---